CMS Virtual Data Overview

Koen Holtman

Caltech/CMS

GriPhyN all-hands meeting, Marina del Rey
April 9, 2001

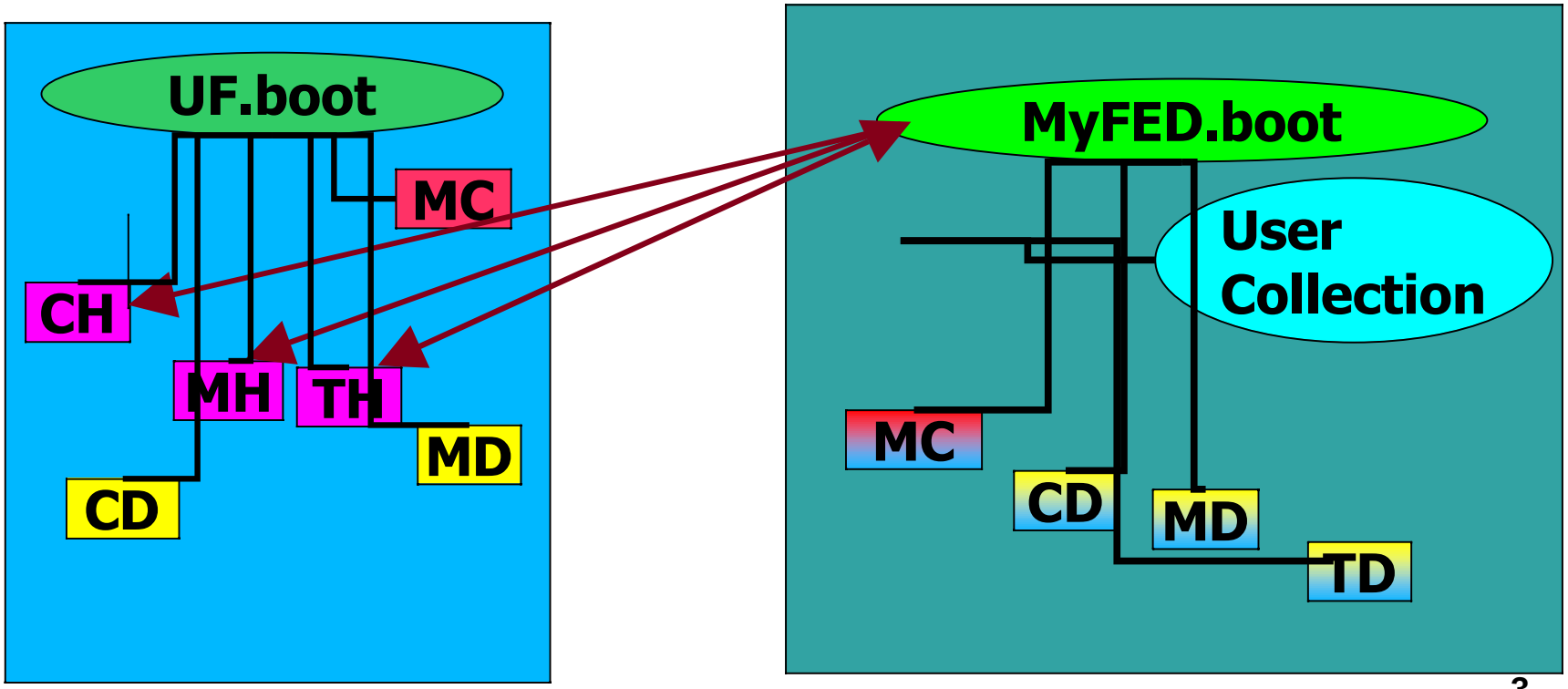
Introduction

- CMS:
  - LHC physics experiment at CERN
  - Participates in **GriPhyN**, **PPDG**, EU DataGrid, ...
  - 2006: Start data taking
    - order of 200 sites, 10,000 CPUs, 5,000 TB
    - order of 500 physicists doing interactive analysis
  - Now: Large-scale simulations of CMS physics
    - order of 10 sites, 1,000 CPUs, 10 TB
    - order of 10 production managers running simulations

- 1995-2000: RD45 project (A Persistent Object Manager for HEP)
- 1997-2000: GIOD project (Globally Interconnected Object Databases)
- 1998- ... : MONARC project (Models of Networked Analysis at Regional Centres for LHC Experiments)

Current CMS data situation

- CMS currently has seamless object-level navigation in the local area
  - Supports private views, deep and shallow copies
  - Uses Objectivity/DB capabilities that do not scale to wide area
    - Therefore we need **Grid** to act as wide-area glue



CMS requirements status

- With respect to all the grid projects (GriPhyN, EU DataGrid, PPDG, ...
- Requirements:
  - CMS internal milestones are already being defined to rely on Grid project deliverables
  - Long term requirements picture for 2006 CMS Grid system:
    - CMS consensus emerging, documents nearing completion
    - (We reserve the right to change our mind later)
  - Shorter term needs:
    - Nothing much documented. Driving factors are:
    - The need to get more automation as the scale of our computing grows -- production managers already overloaded
    - The need to support data access by many more physicists
    - The need to grow ***de facto standard*** solutions for 2006
      - But can't decide on the technology roadmap on our own!

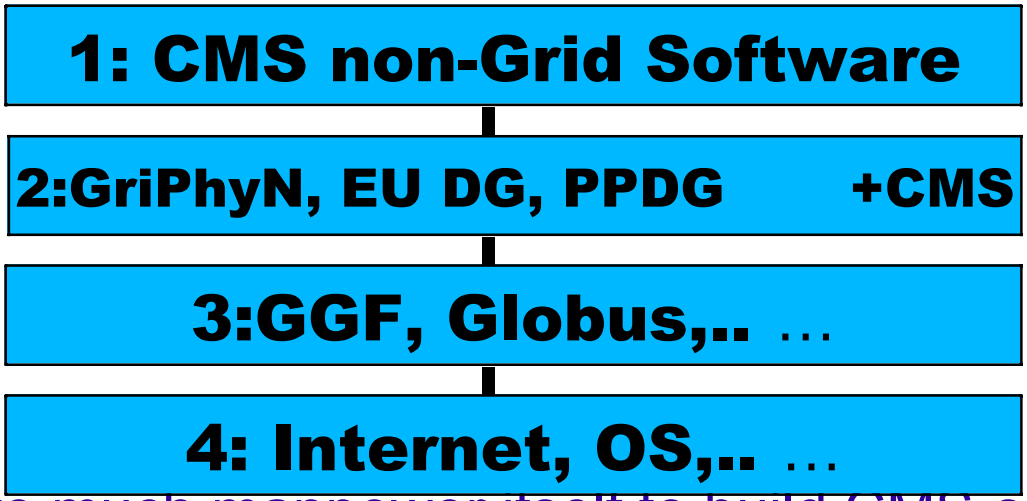
Recent developments

- Many recent discussions (Amsterdam (4-10 March) etc)
  - Picture emerging that there is a crucial need for coordination and compatibility between grid projects
    - LHC data grid problem is simply too large to solve independently; we need to end up with ***one consistent set*** of Grid software
  - Also there is a need for the experiments to coordinate in finding common requirements, adopting common solutions
    - EU DataGrid project now ramping up effort to find common requirements - interaction with the US
- Strategy is:
  - Create one set of (CMS) requirements for all grid projects
  - Participate in discussions between projects; divide the work

2006 CMS grid requirements

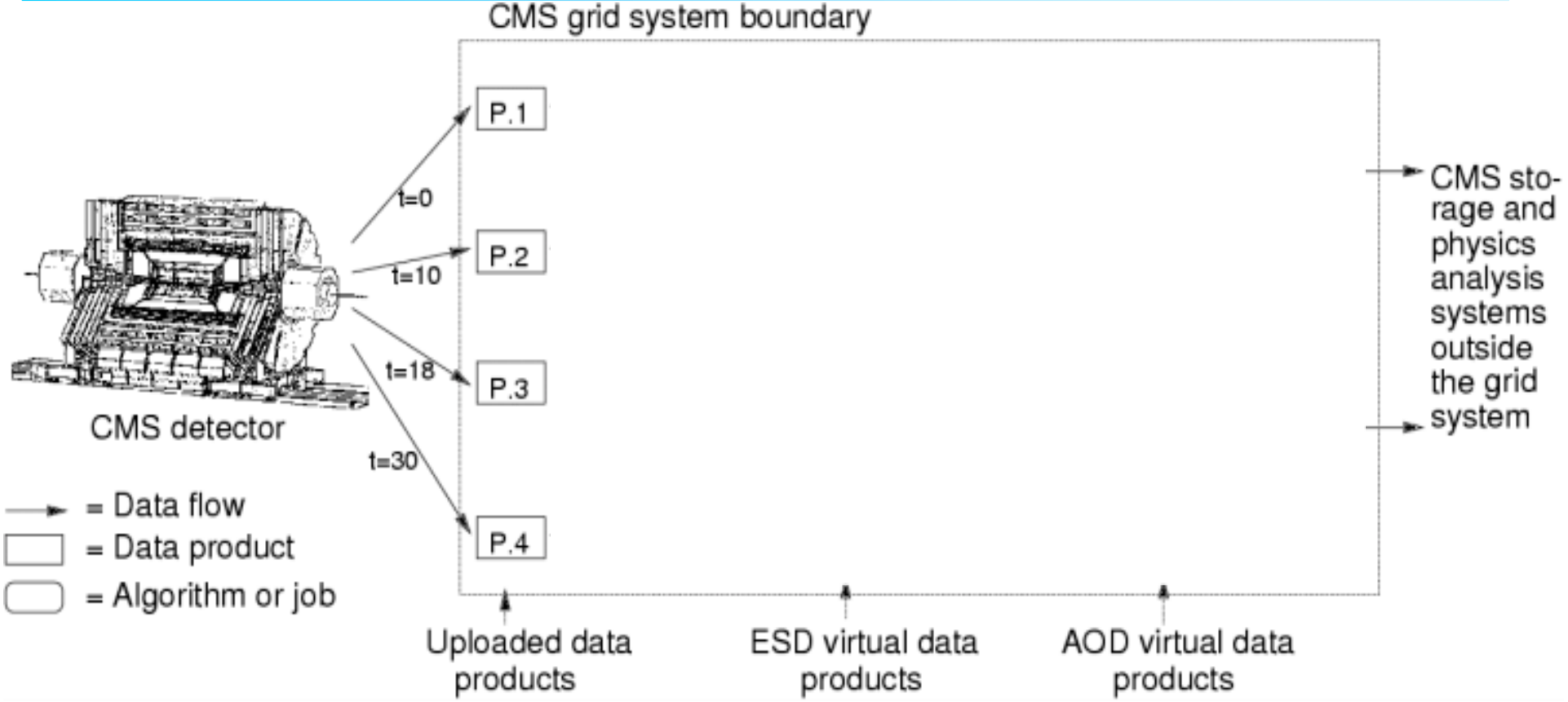
- Specification strategy: describe a baseline for the 2006 production data grid system operated by CMS
  - Emphasize interfaces between box 1 and 2
  - Specify size and placement of hardware, data volumes, …
  - Specify numeric characteristics for interface calls
- From this it follows what the Grid projects should build/research

Component Classification By producer



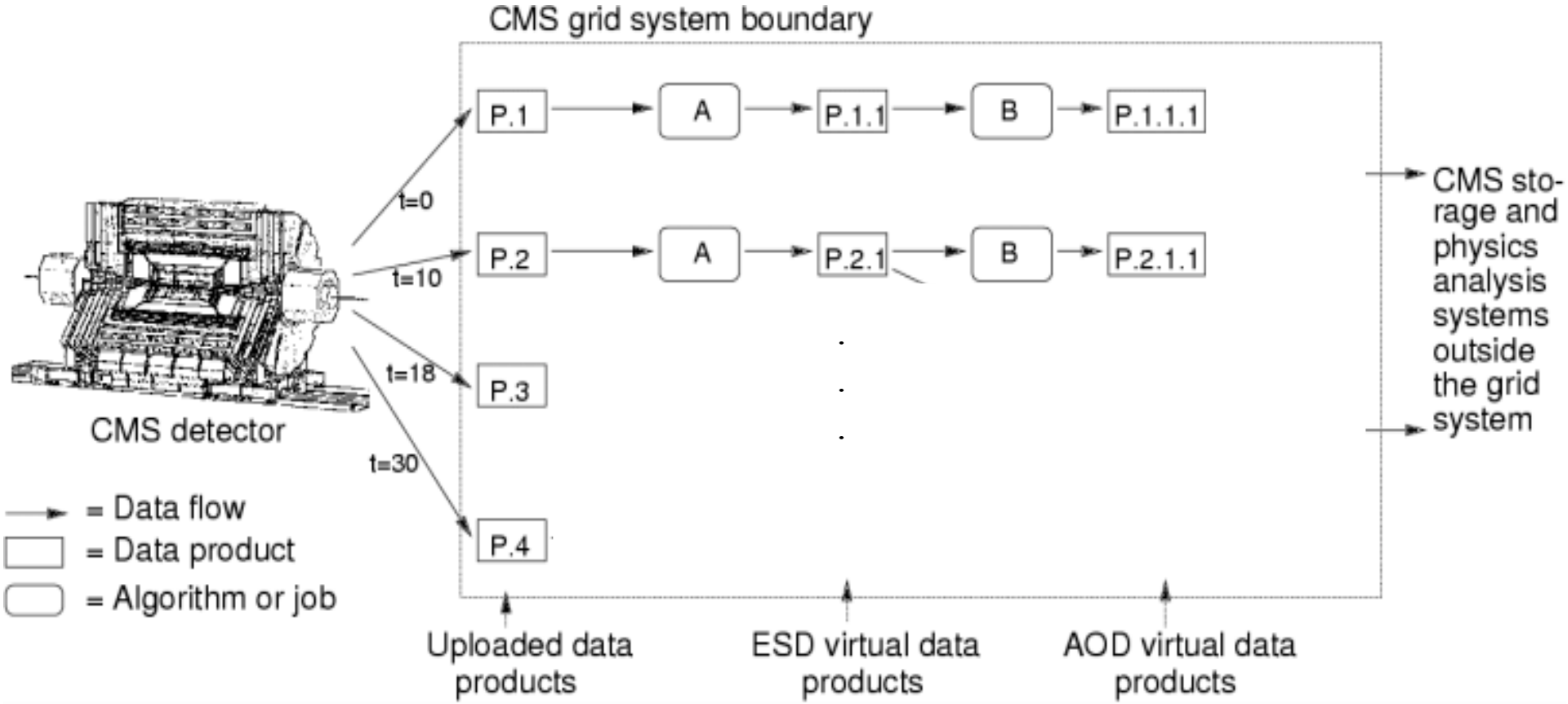
- CMS does not have much manpower itself to build CMS-specific grid components in box 2, rather it is looking to collaborate on joint HEP-specific and `data intensive science-specific' components

Data Model (simplified) 1



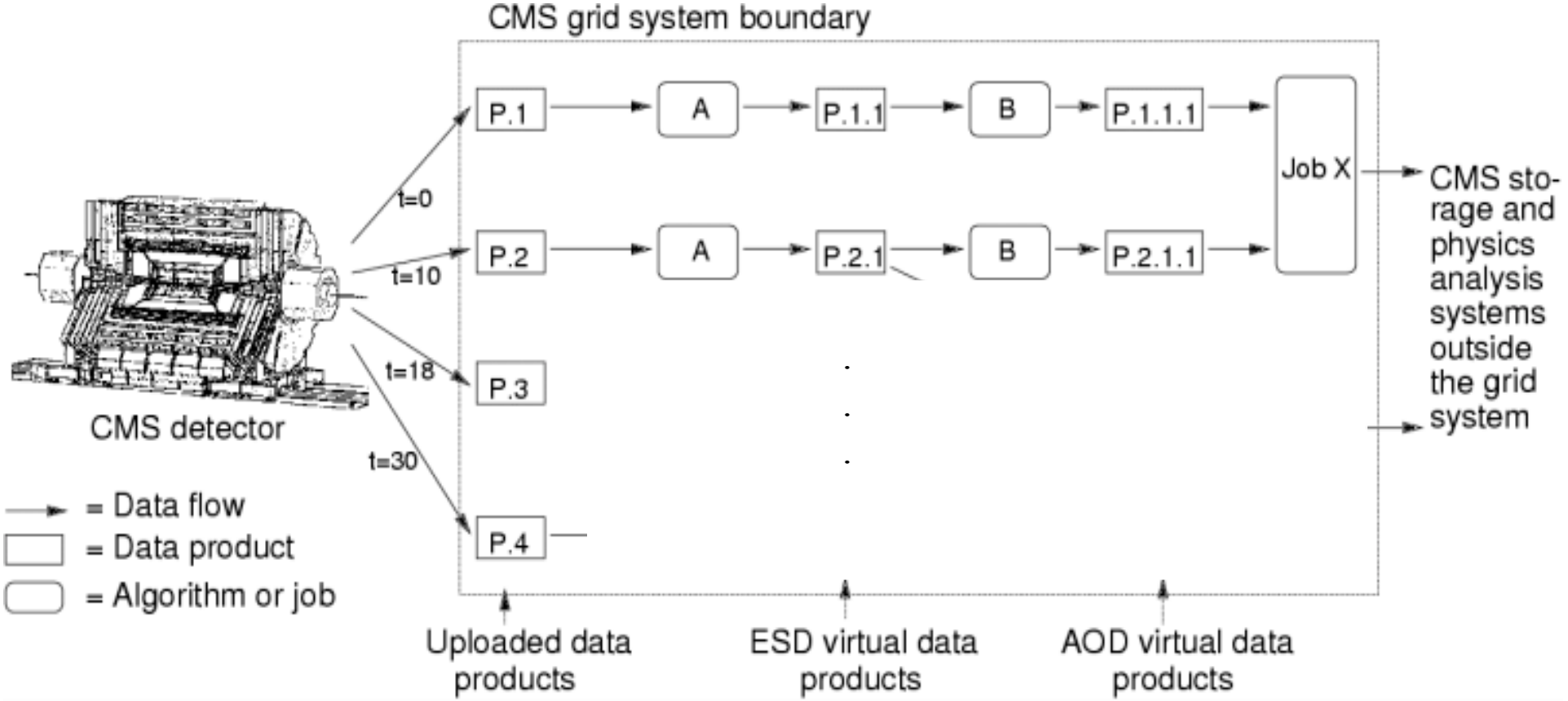
- Data processing chain, but under continuous development
- Some data products are uploaded

Data Model (simplified) 2



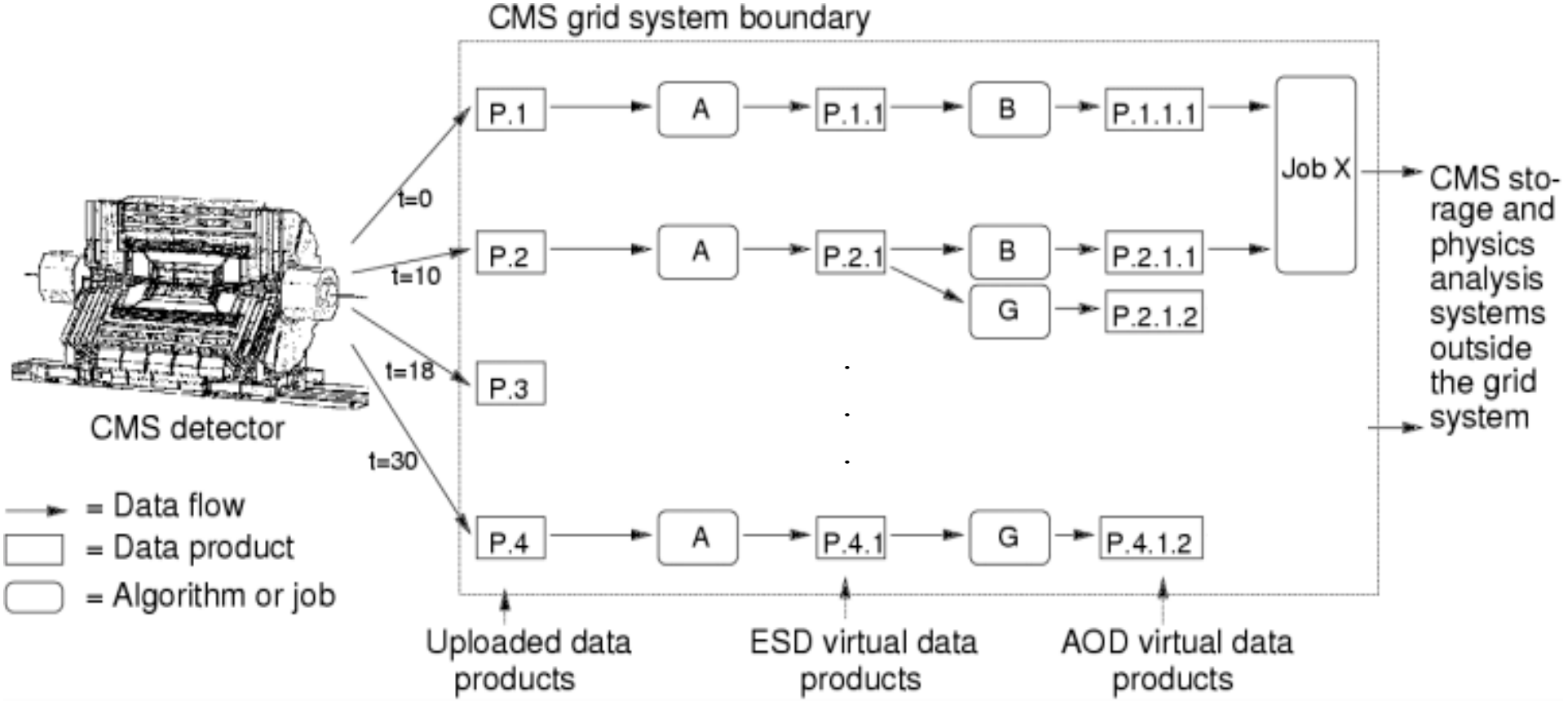
- Algorithms A and B are registered, these define new virtual data products

Data Model (simplified) 3



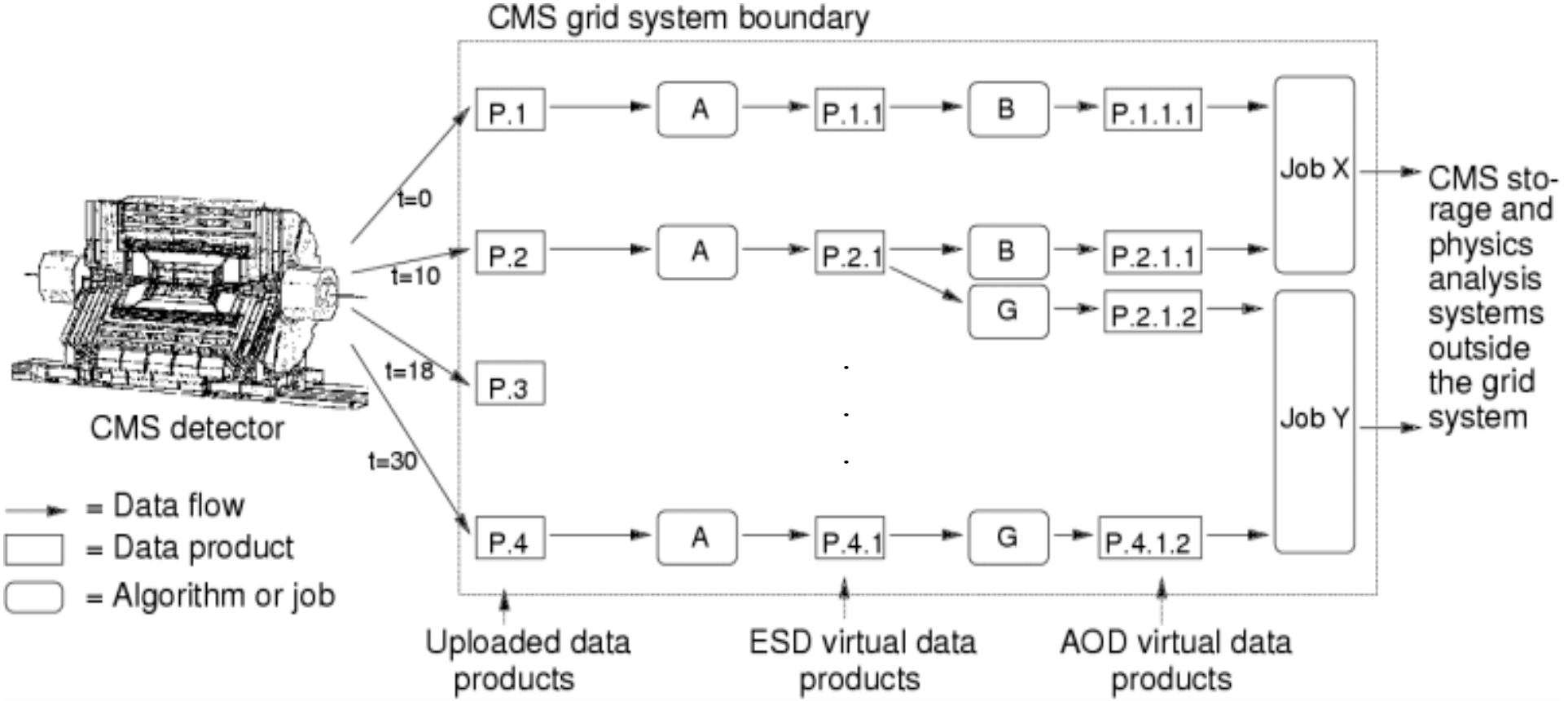
- Job X can now be run to analyze some virtual data products

Data Model (simplified) 4



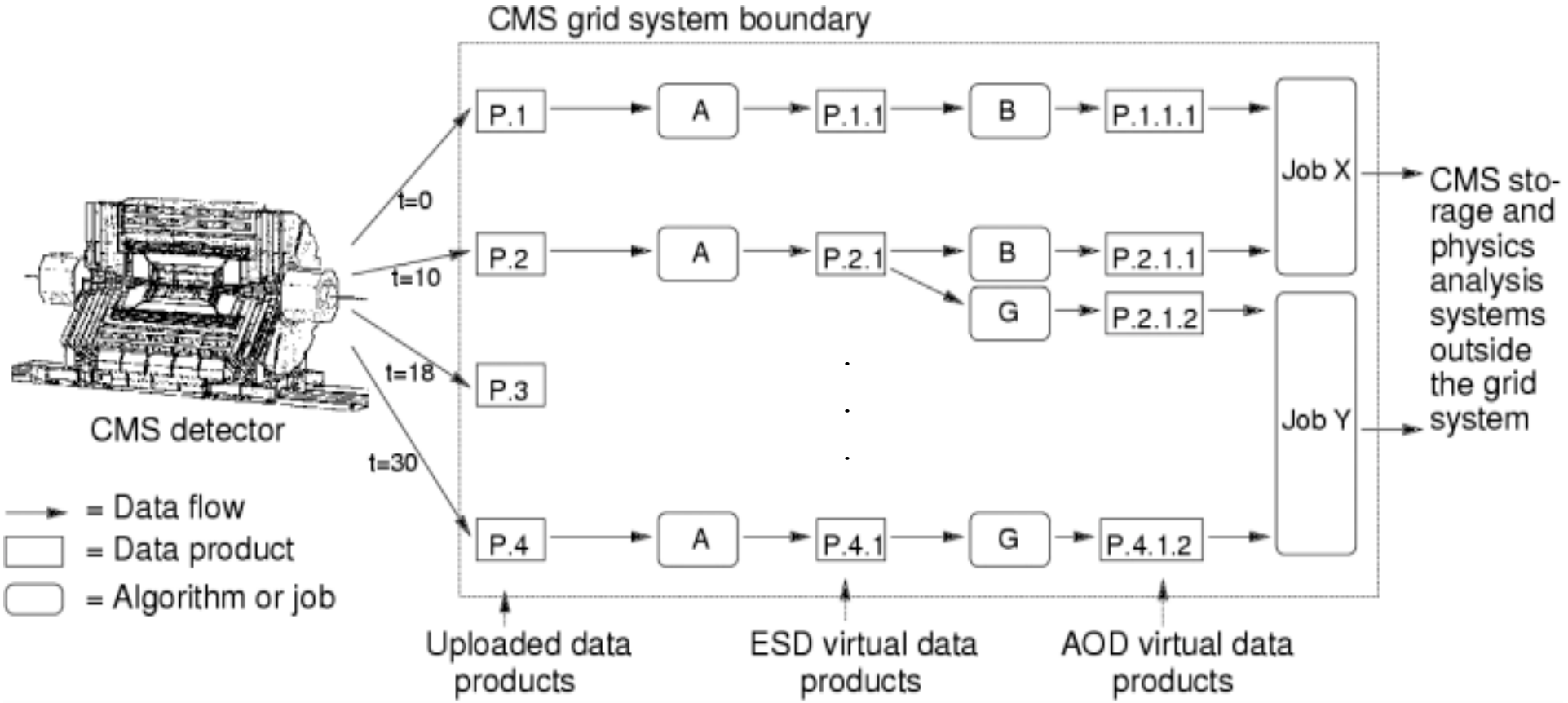
- On the basis of the job X output,
  a new algorithm G is defined and registered,
  this creates new virtual data products

Data Model (simplified) 5



- Data analysis is a cyclic effort of defining and analyzing products
- Uploaded and virtual data products are read-only!
- Virtual data products are small (1 KB -1 MB , <<files)
- >> 10^15 defined, ~10^10 requested in a year
- 50 -10,000 jobs per day

Job Model (simplified)



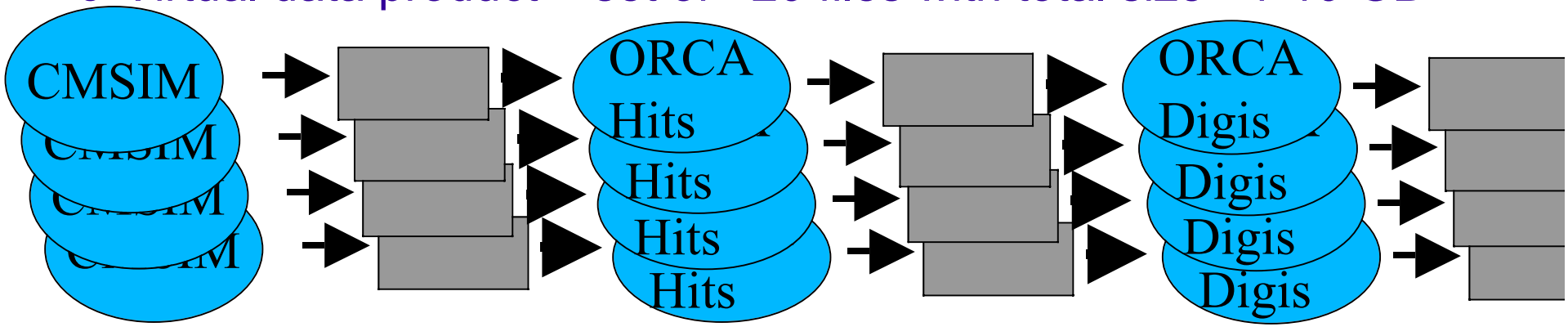
- Jobs are the way by which physicists get work done
- Job = Data product request set (by product name) + Code
- Grid needs to execute code and feed product values to it
- Code is parallel, distributed program
- Job output goes outside the grid

Main research challenges

- Main challenges to suport long-term 2006 CMS needs:
  - Scalability (baseline is 31 tier 0-2 sites with 10,000 CPUs)
  - Fault tolerance (do not have much operator manpower!)
  - Support granularity of our data model
    - >> 10^15 products defined, 10^10 requested in a year
    - Fixed pre-determined mapping of 10^15 virtual products to 10^7(?) virtual files ***will not work!***
    - Materialization and access requests will be too sparse with any conceivable fixed product-file mapping
      - So a flexible mapping from products to files is needed (SC2000 demo)
    - Every one of >>10^15 products needs to have independent materialization status
- Overall research challenge: identify common grid components between experiments

Shorter-term ideas 1

- To support shorter term CMS needs and build towards 2006
- To automate CMS production more: GriPhyN involvement to build **top-level layers of CMS production system in terms of virtual data**.
  - Production use in 1-2 years on a few US CMS sites
  - Virtual data product = set of ~20 files with total size ~1-10 GB



  - Production = requests to materialize these products and store them an archive
- GriPhyN implements top-level virtual data catalog, coupled to scheduler that creates jobs and submits to sites, initiates data movement, provides fault tolerance
- Globus, EU DataGrid, PPDG provide job execution services, file

Shorter-term ideas 2

- The support data access by many more physicists: GriPhyN involvement to build a ***fine-granularity data replication tool*** in terms of the CMS data model
  - 2-D data model (events X event related objects)
  - Physicists select some objects for some event set and have these replicated to a more local installation for further interactive analysis
- This activity could drive the R&D on developing **fine-granularity virtual data** services (>>10^15 virtual products!)
- We believe that this 2-D data model would also fit ATLAS, SDSS, maybe even Ligo data
- SC2000 demo...
  We would like to collaborate with others in developing this

Conclusions

- Initial reaction to architecture documents out of GriPhyN and EU Datagrid:
  - Need for compatibility and dividing the work between the grid projects is obvious
  - Lot of wotk to be done in defining higher-level grid components that encapsulate common experiment needs
- The requirements process is ongoing
  - Current draft document `CMS virtual data requirements'' available from http://home.cern.ch/kholtman
- Inside GriPhyN, we need to discuss and plan
  - CMS milestones related to using GriPhyN products
  - Research directions and relation to other Grid projects
    - Compared to other projects, we see GriPhyN having a heavier emphasis on longer-term needs and developing new grid services, rather than productising existing grid services
  - Testbeds and short-term prototypes
  - Software deliverables (virtual data toolkit components)